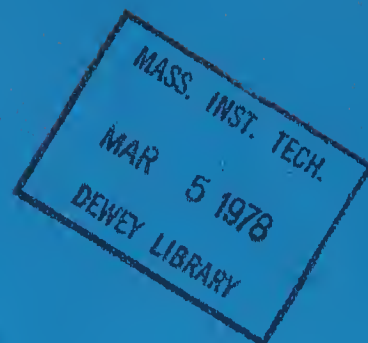


**working paper
department
of economics**



OPTIMAL SEARCH FOR THE BEST ALTERNATIVE

Martin L. Weitzman

Number 214

January 1978

**massachusetts
institute of
technology**
**50 memorial drive
cambridge, mass. 02139**




OPTIMAL SEARCH FOR THE BEST ALTERNATIVE

Martin L. Weitzman

Number 214

January 1978

For their helpful comments, I would like to thank K. Burdett, M. Manove,
M. Obstfeld, and M. Rothschild.



Digitized by the Internet Archive
in 2011 with funding from
Boston Library Consortium Member Libraries

<http://www.archive.org/details/optimalsearchfor00weit>

Optimal Search for the Best Alternative

Summary

This paper completely characterizes the solution to the problem of searching for the best outcome from alternative sources with different characteristics. The optimal strategy is an elementary reservation price rule, where the reservation prices are easy to calculate and have an intuitive economic interpretation.

Introduction

A broad class of economic search problems can be cast in the following form. There are a number of different opportunities or sources, each yielding an unknown reward. The uncertainty about the reward from a source can be eliminated, at a fee, by searching or sampling. Each source has its own: independent probability distribution for the reward; search cost; search time. Sources are sampled sequentially, in whatever order is desired. When it has been decided to stop searching, only one opportunity is accepted, the maximum sampled reward. Under this formulation, what sequential search strategy maximizes expected present discounted value?

A powerful solution concept applies to the above model. Each source is assigned a reservation price -- an invariant critical number analogous to an internal rate of return. The selection rule is to search next that unsampled source with highest reservation price. The stopping rule is to terminate search whenever the maximum sampled reward is above the reservation price of every unsampled source. This simple characterization of an optimal policy is the basic result of the present paper. Its fundamental properties are derived and interpreted.

A Dialogue

The research department of a certain large organization has been assigned the task of finding a new and better way to produce ergs. Two alternative technologies are being considered, the benefits of which are uncertain and would not be known until development work is actually completed. It is estimated that a production process based on the so-called alpha technology could yield a total dollar savings of \$100 with probability $\frac{1}{2}$ and of \$70 with probability $\frac{1}{2}$. The alternative omega process with probability $\frac{1}{5}$ might deliver a possible savings of \$200 but it would not offer any improvement with probability $\frac{4}{5}$. R&D, which must be done to remove the uncertainty, costs \$15 for the alpha process and \$25 for omega. Table 1 summarizes the relevant information.

project	α		Ω	
cost	\$15		\$25	
reward	\$100	\$70	\$200	\$0
probability	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{5}$	$\frac{4}{5}$

Table 1

Department analyst R.D. Findingst is not quite sure what to recommend. Either project looks economical to Findingst, since both are positive by the expected net benefit criterion he remembers from somewhere. But which alternative should be researched first? R.D. doesn't recall how to answer this sort of question. By any criterion he has ever heard of, alpha surely looks better. It has lower cost, higher expected reward, less variance. Findingst senses that alpha must come first, but he is not sure why. Before making a final recommendation he wishes to consult with his old friend, the famous expert on the economics of advanced erg technology, Pandora Erda.

"Pandora," R.D. impores after explaining the situation, "should alpha or omega be developed first? You know the Operations Research guys -- if I guess wrong they'll never let me forget it."

"Don't worry," assures Pandora, "we can figure this out. As you already know, researching either project is better than researching no project. Suppose alpha is developed first. If the payoff turns out to be \$70, you would then want to develop omega, because the expected value of that project would then be

$$-\$25 + \frac{1}{5} (\$200) + \frac{4}{5} (\$70) = \$71$$

which is greater than the value at that point of not developing omega, \$70. However, since

$$-\$25 + \frac{1}{5} (\$200) + \frac{4}{5} (\$100) = \$95 < \$100 ,$$

it would not be economical to develop omega if alpha had a \$100 payoff.

"So the expected value of an optimal policy beginning with developing alpha is

$$-\$15 + \frac{1}{2} (\$100) + \frac{1}{2} (-\$25 + \frac{4}{5} (\$70) + \frac{1}{5} (\$200)) = \$70 \frac{1}{2} .$$

"A similar calculation shows the expected value of an optimal policy which starts by developing omega is

$$-\$25 + \frac{1}{5} (\$200) + \frac{4}{5} (-\$15 + \frac{1}{2} (\$100) + \frac{1}{2} (\$70)) = \$71 .$$

"Hey, it's better to start with omega!"

"What?" exclaims a stunned R.D. "But alpha looks so much more economical. Its net expected value is

$$-\$15 + \frac{1}{2} (\$100) + \frac{1}{2} (\$70) = \$70$$

compared with omega yielding a mere

$$-\$25 + \frac{1}{5} (\$200) + \frac{4}{5} (\$0) = \$15 .$$

"If we had to pick only one project we'd certainly select alpha. It's worth more because we'd be a lot worse off without it than without omega. So how come your conclusion is different?"

"That's just it," muses Pandora, "there is a difference between the worth of a project and the order in which it should be developed. Even though in some sense omega is worth less, the optimal sequential strategy is to develop it first. That way you get a shot at the \$200 right away, and if you fail there is always alpha to fall back on. Hmm. I wonder whether there's a simple selection rule hidden somewhere in all this. If you had to make explicit comparisons among project sequences what a mess that would be in the general case."

"What general case?" snaps R.D. "Wait till the O.R. crew hears what I have to say about analyzing this case!"

Pandora's Problem

There are n closed boxes at the beginning of our scenario. Box i , $1 \leq i \leq n$, contains a potential reward of x_i with probability distribution function $F_i(x_i)$, independent of the other rewards. It costs c_i to open box i and learn its contents, which become known only after a time lag of t_i . Instantaneous learning is the special case $t_i = 0$.

An initial amount x_0 is available, representing a fallback reward that could always be collected if no sampling were undertaken or if every sampled reward happened to be less than x_0 . In many applications it is natural to set $x_0 = 0$. All costs and benefits are converted to present values by the discount rate r .

At each stage Pandora must decide whether or not to open a box. If she chooses to stop searching, Pandora collects at that time the maximum reward she has thus far uncovered. Should Pandora wish to continue sampling, she must select the next box to be opened, pay at that time the fee for opening it, and wait for the outcome. Then will come the next decision stage. Note a characteristic asymmetry: the sum of search costs is paid during search, whereas the maximum reward is collected after search has been terminated.

Pandora worships maximized expected present discounted value. She needs to know what she should do to be consistent with this fundamental conviction. Pandora wants a sequential decision rule that will tell her at each stage whether or not to continue searching, and if so, which box to open next.

Pandora's problem can be formally posed in dynamic programming format. Let the collection of n boxes, denoted I , be partitioned into any set S of sampled boxes and its complement \bar{S} of closed boxes. That is,

$$S \cup \bar{S} = I, S \cap \bar{S} = \phi$$

where

$$I = \{1, 2, \dots, n\}.$$

The variable y will represent the maximum sampled reward (from the opened boxes and the initial fallback reward)

$$y = \max_{i \in S \cup \{0\}} x_i. \quad (1)$$

It is intuitively obvious (and easily verified) that all relevant information about the previously opened boxes is summarized by y ; knowing the individual values of x_i for $i \in S \cup \{0\}$ is superfluous to making a correct decision because all probability distributions are independent.

The state of the system at any time is given by the statistic (\bar{S}, y) . Define $\Psi(\bar{S}, y)$ as the expected present discounted value of following an optimal policy from this time on when the set of closed boxes is \bar{S} and the maximum sampled reward is y .

For each subset \bar{S} of I and every y , the state valuation functions must satisfy the fundamental recursive relation

$$\Psi(\bar{S}, y) = \max \left\{ y, \max_{i \in \bar{S}} \left\{ -c_i + \beta_i \left[\Psi(\bar{S} - \{i\}, y) \int_{-\infty}^y dF_i(x_i) + \int_y^{\infty} \Psi(\bar{S} - \{i\}, x_i) dF_i(x_i) \right] \right\} \right\} \quad (2)$$

where

$$\Psi(\emptyset, x) = x \quad (3)$$

$$\beta_i = e^{-rt_i} \quad (4)$$

Equation (2) is just the principle of optimality for dynamic programming. At stage (\bar{S}, y) Pandora could terminate search, collecting reward y . Or, she might open box i , for each $i \in \bar{S}$, which results in expected discounted net gain

$$-c_i + \beta_i \left[\Psi(\bar{S} - \{i\}, y) \int_{-\infty}^y dF_i(x_i) + \int_y^{\infty} \Psi(\bar{S} - \{i\}, x_i) dF_i(x_i) \right].$$

The value of an optimal policy at (\bar{S}, y) is the maximum of these alternatives.

In principle, the state valuation functions $\{\Psi(\bar{S}, y)\}$ could be recursively built up by systematic induction on the number of closed boxes. Using (2), (3), state valuation functions could be constructed first for all sets consisting of one closed box, then for all sets of two boxes, of three, four, etc. The actual computation is likely to be a combinatoric task of unwieldy proportions unless the number of boxes is very small.

At any stage (\bar{S}, y) , Pandora's optimal decision is that policy which maximizes the right hand side of (2). If two or more policies tie, it makes no difference how the tie is broken. Note that although an optimal strategy is implicitly contained in equation (2), the form of that strategy is little more than a complete enumeration of what to do in all possible situations.

The search literature has dealt extensively with the situation where, in effect, all boxes are identical. For this special case the issue of choosing which box to open does not arise. The essential question is when to stop. The answer is: search continues until a reward greater than some "reservation price" is discovered. The reservation price for sampling with recall is that hypothetical cutoff value of the maximum reward which would make it just equal to the expected net gain of opening exactly one more box.¹

The contribution of the present paper is to show that with alternative search opportunities the optimal policy is a straightforward analogue of the above idea. Each (different) box is assigned a (different) reservation price *calculated by the same formula as before,* which now serves as a basis for the optimal stopping and selection rule. The reservation price of a box determines its ordinal ranking, prescribing when it should be opened relative to the other boxes. Thus, all the advantages of a simple rate of return criterion apply in a search context.

Some Applications

The formulation presented in this paper is general enough to cover, at a high level of abstraction, economic search models from a variety of settings.

¹This stopping rule is well known and appears in many places. See, for example, Lippman and McCall [1976] or Landsberger and Peled [1977].

Take for example the standard job search model with wage offers retained. The current framework allows the situation where the job searcher may choose to sample from various firms having different characteristics. The lump sum reward is most appropriately interpreted as the discounted present value of all future wages. Search costs, which presumably include a psychic component, are net of any side compensation (unemployment benefits or wages from a currently held job). The possibility of reaccepting current work while searching on-the-job is accommodated by making the fallback reward x_0 equal to the present discounted value of the current wage. Other modifications are also possible.

Searching for the lowest price on some commodity available from different stores is also an example amenable to the analysis developed in this paper. Let the good have some intrinsic utility measured in dollar terms. The reward available from a store is the difference between the utility of the commodity and its price. Search costs should include the opportunity loss of forgoing the item in question while search continues, as well as the more orthodox cost of visiting a store to obtain a price quotation. The option of not buying the good at all can be represented by setting the fallback reward equal to the (dis)utility of henceforth doing without the item altogether, which can be normalized to zero.

Another area of application concerns the optimal sequential research strategy for developing various uncertain technologies to meet the same or a similar purpose. The reward is the potential cost saving of the new technology, unknown until after it has been developed. Search fees are research and development expenditures. Search time is the anticipated length of the R&D process. The option of choosing to continue with the current known technology is represented by having a zero fallback reward.

There are other possible interpretations of the model, but it would be tedious to go on listing them.

The Optimal Strategy

Any closed box is characterized by: a fee for opening it; a time lag for discovering its contents; a probability distribution for the reward it contains. Suppose all this information must somehow be compressed into a single index number, a kind of internal rate of return. One heuristic procedure would be to evaluate the intrinsic search value of a closed box by assigning it the hypothetical reward of that opened box to which it is in some sense equivalent.

Suppose for the moment there are just two boxes. One is the closed box i . The other is an already opened hypothetical box offering reward z_i . If the searcher elects not to open box i , she receives the "sure thing"

$$z_i . \quad (5)$$

If she opens box i , the searcher can expect a net benefit

$$-c_i + \beta_i \left[z_i \int_{-\infty}^{z_i} dF_i(x_i) + \int_{z_i}^{\infty} x_i dF_i(x_i) \right] . \quad (6)$$

The closed and opened boxes are heuristically "equivalent" if the searcher is just indifferent between opening box i and not opening it. This will occur if (5) and (6) are equal to each other, a condition which can be rewritten

$$c_i = \beta_i \int_{z_i}^{\infty} (x_i - z_i) dF_i(x_i) - (1 - \beta_i) z_i . \quad (7)$$

The critical number z_i which satisfies (7) is called the reservation price of box i .

Although the definition (7) has been motivated by heuristic considerations, it turns out there is a rigorous sense in which all relevant information about box i is summarized by its reservation price z_i .

The following decision strategy, called Pandora's rule, completely characterizes an optimal policy.

Selection Rule: If a box is to be opened, it should be that closed box with highest reservation price.

Stopping Rule: Terminate search whenever the maximum sampled reward exceeds the reservation price of every closed box.

What is remarkable about this rule is that the entire structure of an optimal policy has been reduced to a simple statement about reservation prices. Furthermore, the reservation price of each box is calculated by equating a hypothetical gain of stopping (5) not with the full gain of opening the box and continuing on in an optimal manner, but rather with the myopic gain of opening the box and terminating (6). In other words, the reservation price of a box depends only on the properties of that box and is independent of all other search opportunities.

Note that if Pandora samples from n identical boxes, the optimal policy is to continue search until she uncovers a reward greater than the common reservation price of each box. In this special case, it is comparatively simple to prove optimality.

The proof of Pandora's rule, which is quite technical, is relegated to the final section.

From (7), the reservation price of a box is completely insensitive to the probability distribution of rewards at the lower end of the tail. Any rearrangement of the probability mass located below z_i leaves z_i unaltered. It is important to understand this feature. Considering that a box could be opened at any time, the only rationale for opening it now is the possibility of terminating further search by drawing a relatively high reward. That is why the lower end of its reward distribution is irrelevant to the order in which box i should be sampled even though it may well influence the value of an optimal policy by altering the likelihood that x_i will end up being the largest reward drawn.

On the other hand, as rewards become more dispersed at the upper end of the distribution, the reservation price increases and so does the net benefit of search. Other things being equal, it is optimal to sample first from distributions which are more spread out or riskier in hopes of striking it rich early and ending the search. This is a major result of the present paper. Low-probability high-payoff situations should be prime candidates for early investigation even though they may have a smaller chance of ending up as the source ultimately yielding the maximum reward when search ends.

The standard comparative statics exercises performed on (7) yield anticipated results. Reservation price decreases with: greater search cost, increased search time, or a higher interest rate. Moving the probability mass of rewards to the right (i.e., changing the distribution function $F_i(x_i)$ to $G_i(x_i) \leq F_i(x_i)$) makes z_i larger. Thus, although there is no necessary connection between the mean reward and the reservation price, there is a well-defined sense in which higher rewards increase the reservation price. Similarly, performing a mean preserving spread

on the distribution function $F_i(x_i)$ makes z_i bigger. In this sense a riskier distribution of rewards implies a higher reservation price.²

Note that acceptance levels decline with the duration of search, as the best opportunities are sampled first and the poorer ones later.

Because it is so easy to calculate reservation prices, sensitivity analysis is made especially simple. The effect on project ranking (and hence on an optimal policy) of changing such parameters as search costs, the probability distribution of rewards, search time, or the interest rate is easily determined. It is also easy to say how an optimal search strategy changes when certain opportunities are added to or deleted from the list of prospective candidates.

An Example

To illustrate the nature of the solution concept and indicate what it depends on, an explicit example is calculated in an interesting special case.

Suppose for each i , box i contains one of two outcomes: either zero reward ("failure") with probability $1-p_i$, or positive reward R_i ("success") with probability p_i . To keep things simple there is no discounting ($\beta_i = 1$) and the expected net gain $p_i R_i - c_i$ is positive.

Applying (7) to this special case yields the closed form expression

$$z_i = \frac{p_i R_i - c_i}{p_i} . \quad (8)$$

The reservation price of a box is the expected net gain divided by the probability of success. For the same expected net gain, that box is opened first which offers a smaller probability of success.

²This feature has been analyzed by Kohn and Shavell [1974] for the case of identical boxes.

on the distribution function $F_i(x_i)$ makes z_i bigger. In this sense a riskier distribution of rewards implies a higher reservation price.²

Note that acceptance levels decline with the duration of search, as the best opportunities are sampled first and the poorer ones later.

Because it is so easy to calculate reservation prices, sensitivity analysis is made especially simple. The effect on project ranking (and hence on an optimal policy) of changing such parameters as search costs, the probability distribution of rewards, search time, or the interest rate is easily determined. It is also easy to say how an optimal search strategy changes when certain opportunities are added to or deleted from the list of prospective candidates.

An Example

To illustrate the nature of the solution concept and indicate what it depends on, an explicit example is calculated in an interesting special case.

Suppose for each i , box i contains one of two outcomes: either zero reward ("failure") with probability $1-p_i$, or positive reward R_i ("success") with probability p_i . To keep things simple there is no discounting ($\beta_i = 1$) and the expected net gain $p_i R_i - c_i$ is positive.

Applying (7) to this special case yields the closed form expression

$$z_i = \frac{p_i R_i - c_i}{p_i} . \quad (8)$$

The reservation price of a box is the expected net gain divided by the probability of success. For the same expected net gain, that box is opened first which offers a smaller probability of success.

²This feature has been analyzed by Kohn and Shavell [1974] for the case of identical boxes.

After ranking boxes to be opened in order of decreasing z_1 , the searcher moves down the list until a success is encountered. At that point search ends because $R_1 > z_1$.

Suppose, as a further restriction, search is for the same object, like a new product with certain well-defined characteristics. Then all rewards R_1 are identical. In that case (8) reduces to

$$z_1 = R - \frac{c_1}{p_1}.$$

The best opportunity is the one offering the highest probability of success per dollar of search cost.

Some Limitations

The purpose of the model formulated in this paper is to sharply characterize optimal search among alternative sources with different characteristics. Naturally certain "other" aspects of the optimal search problem have been abstracted away.

Many of the underlying assumptions of the present formulation are unrealistic. There has been no provision made for: adaptive learning about correlated probability distributions; pay-as-you-go research (with the possibility of backing out of a project if prospects start looking unfavorable); parallel search activity; risk aversion; incomplete or no recall; collecting some reward before search is terminated; randomly generated new opportunities; a binding time horizon; uncertain search costs or search time; etc.³ Yet the model as a whole captures enough essential aspects of reality that it should be useful in providing project rankings which

³ Some of these topics have been treated in the literature, most typically for the symmetric case where all boxes are identical. See the bibliography cited at the end of this paper.

might serve as a rough planning guide of sorts, a kind of pre-investment screening device, or a reference point for the numerical analysis of a more comprehensive dynamic programming type formulation.

The fact that it is possible to explicitly construct an optimal solution makes the problem analyzed here a natural preliminary to more general formulations. And the present model may even be a reasonable description of some situations.

That such an elementary decision strategy as Pandora's rule is optimal depends more crucially than might be supposed on the simplifying assumptions of the model. There does not seem to be available a sharp characterization of an optimal solution when certain features of the present model are changed. Pandora's rule does not readily generalize.

For example, Pandora's rule is not optimal if she may only open m of the n boxes available to her. An example of this was provided in the second section, where α was preferable when only one opportunity could be searched, whereas ω was the better starting choice with the possibility of sequential sampling from both sources. In the general case $n > m \geq 2$, an involved permutational exercise would be required to determine which box should be sampled next. The choice would depend in a complicated way on the maximum reward that has already been drawn and the properties of all the boxes which remain unopened.

If reward distributions were not independent, the optimal search strategy could be very intricate. When a box is opened, the searcher would learn not only about its contents, but also about the reward distributions of alternative boxes.⁴ It appears plausible that other things

⁴Rothschild [1974] contains an illuminating analysis of adaptive search policies for the case of identical boxes.

being equal it would be better to open a box whose reward is highly correlated with other rewards because this adds a positive informational externality. But translating such an effect into a simple search rule seems difficult except in the most elementary cases.

Parallel search efforts and pay-as-you-go research with the option of backing out are important features of the R&D scene omitted from the current formulation. They seem to be very hard to model well.⁵ Perhaps it is wishful thinking, but my feeling is that the results of this paper might still constitute a useful guide here. Even though diluted by more complicated and realistic considerations, some prescription resembling Pandora's rule should remain a prominent feature of any optimal sequential search policy.

If some part of its reward can be collected from a research project before the sequential search procedure as a whole is terminated, that could negate Pandora's rule in extreme cases. It might be optimal to start off with a cheap low-risk research project which promises to supply modest benefits throughout the period of sequential search for the best alternative. Such a project is unlikely to be chosen at the end of the search, but it is developed at the beginning because it can provide a stream of interim rewards while the results of further sampling are awaited.

The cases of sampling without recall, risk aversion, and randomly generated new opportunities have been treated to some extent in the literature.⁶ An optimal policy is typically complicated, especially when there are different kinds of boxes. What to do next will depend in an intricate way on past results and future possibilities.

⁵See, for example, Marshak et al. [1967].

⁶See especially Salop [1973] who treats thoroughly the no recall case. The other two cases are briefly surveyed for a situation with identical boxes in Lippman and McCall [1976].

If search costs or search times are randomly distributed independently of everything else, no changes in formulation are necessary so long as c_i and β_i are interpreted as mean values.

Proof of Optimality

First it must be shown that the formula for determining the reservation price of a box is well defined.

Let

$$H_i(z) = \beta_i \int_z^{\infty} (x_i - z) dF_i(x_i) - (1 - \beta_i)z . \quad (9)$$

It is easily verified that the function $H_i(z)$ is continuous and monotonic. By taking the appropriate limits, $H_i(-\infty) = \infty$, $H_i(\infty) = -\infty$ ($=0$ if $\beta_i=1$). Thus, there exists a solution z_i to the equation

$$c_i = H_i(z_i)$$

which is unique.

The proof of the main proposition is by induction on the number of closed boxes. Suppose Pandora's rule is optimal with m closed boxes remaining and any value of y (representing the maximum return from the previously opened boxes and the initial fallback reward). For $m = 1$, the optimality of Pandora's rule is easily demonstrated just by directly applying the definition of reservation price.

Henceforth we will be considering a situation with $m+1$ closed boxes (the set \bar{S}) and any value of y .

Let j be a box with biggest reservation price in the collection of $m+1$ closed boxes

$$j \in \bar{S}$$

$$z_j = \max_{i \in \bar{S}} z_i . \quad (10)$$

If $y \geq z_j$, it is simple to demonstrate the optimality of not opening any boxes. (After one box is opened, by Pandora's rule applied to m closed boxes it will be optimal to stop. Hence the question is whether opening exactly one box is better than not opening any, which is easily answered in the negative.) The stopping criterion of Pandora's rule is thus proved for $m+1$ closed boxes.

If $y < z_j$, it is straightforward to show the nonoptimality of no further search. (Just opening box j and then stopping would yield a higher expected present discounted value.) Thus, at least one box should be opened.

Suppose (by contradiction with Pandora's rule) it is optimal to open box k first, where k is any box in \bar{S} having a lower reservation price than j

$$k \in \bar{S}$$

$$z_k < z_j . \quad (11)$$

If box k is opened first, by the induction assumption on Pandora's rule for m closed boxes there is an exact prescription of what to do in an optimal policy thereafter. Let the expected discounted present value of opening box k and following Pandora's rule thereafter, which is alleged to constitute the best strategy, be B .

Consider the following alternative. Open box j first. Let h be a box with second biggest reservation price in the collection of $m+1$ closed boxes

$$h \in \bar{S} - \{j\}$$

$$z_h = \max_{i \in \bar{S} - \{j\}} z_i . \quad (12)$$

If $x_j \geq z_h$, terminate. Otherwise, open box k next. From then on proceed by Pandora's rule. Let the expected present discounted value of this alternative policy be A .

The rest of the proof, although technical in its details, essentially consists of showing that $A > B$. From this it follows that the originally proposed policy of first opening box k cannot in fact be best. It must be optimal to first open the box with biggest implicit worth in \bar{S} and then, from the induction assumption, to proceed by Pandora's rule for the remaining m closed boxes. But this is just Pandora's selection rule for $m+1$ closed boxes, completing the induction step.

The following notation is employed (Figure 1 may be useful in providing a sort of mnemonic device).

$$\pi_j = \text{prob}(x_j > z_j) \quad w_j = E[x_j | x_j > z_j]$$

$$\pi_k = \text{prob}(x_k > z_j) \quad w_k = E[x_k | x_k > z_j]$$

$$\lambda_j = \text{prob}(z_h < x_j < z_j) \quad v_j = E[x_j | z_h < x_j < z_j] \quad \tilde{v}_j = E[\max(x_j, y) | z_h < x_j < z_j]$$

$$\lambda_k = \text{prob}(z_h < x_k < z_j) \quad v_k = E[x_k | z_h < x_k < z_j] \quad \tilde{v}_k = E[\max(x_k, y) | z_h < x_k < z_j]$$

$$\mu_k = \text{prob}(z_k < x_k < z_h) \quad u_k = E[x_k | z_k < x_k < z_h]$$

variable:	x_k	x_j	x_k	x_j	x_k
probability:	μ_k	λ_j	λ_k	π_j	π_k
value(s):	u_k	v_j	v_k	w_j	w_k
		\tilde{v}_j	\tilde{v}_k		

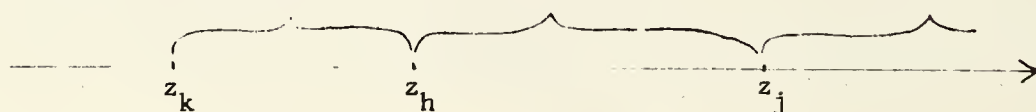


Figure 1

$$d = E[\max(x_j, x_k, y) | z_h < x_j < z_j; z_h < x_k < z_j] \quad (13)$$

$$\Phi = E[\Psi(\bar{S} - \{j\} - \{k\}, \max(x_j, x_k, y)) | x_j < z_h; x_k < z_h] \quad (14)$$

The expected present discounted value of opening box k and then proceeding by Pandora's rule is

$$\begin{aligned} B = & -c_k + \pi_k \beta_k w_k + \lambda_k \beta_k [-c_j + \pi_j \beta_j w_j + \lambda_j \beta_j d + (1 - \pi_j - \lambda_j) \beta_j \tilde{v}_k] \\ & + (1 - \pi_k - \lambda_k) \beta_k [-c_j + \pi_j \beta_j w_j + \lambda_j \beta_j \tilde{v}_j] + (1 - \pi_k - \lambda_k)(1 - \pi_j - \lambda_j) \beta_k \beta_j \Phi. \end{aligned} \quad (15)$$

The proposed alternative is: open box j ; if $x_j \geq z_h$, terminate; if $x_j < z_h$, next open box k and then proceed by Pandora's rule. The expected present discounted value of such a policy is

$$\begin{aligned} A = & -c_j + \pi_j \beta_j w_j + \lambda_j \beta_j \tilde{v}_j + (1 - \pi_j - \lambda_j) \beta_j [-c_k + \pi_k \beta_k w_k + \lambda_k \beta_k \tilde{v}_k] \\ & + (1 - \pi_j - \lambda_j)(1 - \pi_k - \lambda_k) \beta_j \beta_k \Phi. \end{aligned} \quad (16)$$

Subtracting (15) from (16), cancelling some terms and grouping others,

$$\begin{aligned} A - B = & [c_j - \pi_j \beta_j w_j] ((1 - \pi_k) \beta_k - 1) + (c_k - \pi_k \beta_k w_k) (1 - (1 - \pi_j - \lambda_j) \beta_j) \\ & + \lambda_j \beta_j \tilde{v}_j - \lambda_k \beta_k \lambda_j \beta_j d - (1 - \pi_k - \lambda_k) \beta_k \lambda_j \beta_j \tilde{v}_j. \end{aligned} \quad (17)$$

From (7),

$$c_j = \beta_j \pi_j (w_j - z_j) - (1 - \beta_j) z_j, \quad (18)$$

$$c_k = \beta_k [\pi_k (w_k - z_k) + \lambda_k (v_k - z_k) + \mu_k (u_k - z_k)] - (1 - \beta_k) z_k. \quad (19)$$

Substituting in (17) for c_j and c_k from the above expressions yields something that can be manipulated into the form

$$\begin{aligned} A - B = & (z_j - z_k) [(\pi_j \beta_j + 1 - \beta_j)(\pi_k \beta_k + 1 - \beta_k)] + (v_k - z_k) [\lambda_k \beta_k (1 - \beta_j + \pi_j \beta_j)] \\ & + (\tilde{v}_j - z_k) [\lambda_j \beta_j (1 - \beta_k + \pi_k \beta_k)] + (u_k - z_k) [\mu_k \beta_k (1 - \beta_j + \pi_j \beta_j + \lambda_j \beta_j)] \\ & + (\tilde{v}_j + v_k - z_k - d) [\lambda_k \beta_k \lambda_j \beta_j]. \end{aligned} \quad (20)$$

From the definition (13),

$$\begin{aligned}
 d &= z_h + E[\max(\max(x_j, y) - z_h, x_k - z_h) / z_h \leq x_j < z_j; z_h \leq x_k < z_j] \\
 &\leq z_h + E[(\max(x_j, y) - z_h + x_k - z_h) / z_h \leq x_j < z_j; z_h \leq x_k < z_j] \\
 &= \tilde{v}_j + v_k - z_h \\
 &\leq \tilde{v}_j + v_k - z_k .
 \end{aligned} \tag{21}$$

Using the above inequality and the fact that $0 < \beta_j \leq 1$, $0 < \beta_k \leq 1$, every term of expression (20) is seen to be non-negative, with the first term strictly positive. Thus,

$$A > B . \tag{22}$$

This concludes our proof of the form of an optimal policy.

Strictly speaking, we have proved the necessity of Pandora's rule. That rule specifies a unique strategy for each state (except when there is a tie for the maximum reservation price of a closed box, in which case it can be shown, along the lines of the current proof, that how the tie is broken makes no difference to the value of the objective function). Thus, since an optimum policy exists, sufficiency of Pandora's rule has also been demonstrated.

References

Chow, Y.S., Herbert Robbins, and David Siegmund, Great Expectations:

The Theory of Optimal Stopping, Houghton-Mifflin, New York, 1971.

De Groot, Morris H., Optimal Statistical Decisions, McGraw-Hill, New York, 1970.

Kohn, Meir G., and Steven Shavell, "The Theory of Search," Journal of Economic Theory, October 1974, 9, 93-123.

Landsberger, Michael, and Dan Peled, "Duration of Offers and Price Realizations," Journal of Economic Theory, October 1977, 17-37.

Lippman, Steven A., and John J. McCall, "The Economics of Job Search: A Survey," Economic Inquiry, June 1976, 155-189.

Marshak, T., T. Glennon, and R. Summers, Strategy for R&D, Springer-Verlag, 1967.

Rothschild, Michael, "Searching for the Lowest Price When the Distribution of Prices is Unknown," Journal of Political Economy, July-August 1974, 82, 689-711.

Salop, Steven C., "Systematic Job Search and Unemployment," Review of Economic Studies, April 1973, 40, 191-201.

Date Due

<div data-bbox="366 588 496 635">JAN 12 '79</div> <div data-bbox="329 661 461 709">MAR 30 '79</div> <div data-bbox="347 709 478 749">FEB 16 '80</div> <div data-bbox="324 756 434 796">DEC 06 '81</div> <div data-bbox="361 830 557 876">APR 30 1987</div> <div data-bbox="399 937 595 983">APR 30 1987</div> <div data-bbox="315 1058 478 1108">MAR 07 '89</div> <div data-bbox="315 1112 525 1159">MAR 07 '89</div> <div data-bbox="343 1139 525 1189">DEC 13 1992</div> <div data-bbox="343 1219 525 1280">MAY 15 '88</div>		
--	--	--

MIT LIBRARIES



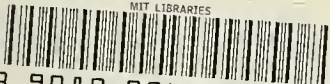
3 9080 004 446 362

MIT LIBRARIES



3 9080 004 415 375

MIT LIBRARIES



3 9080 004 415 383

MIT LIBRARIES



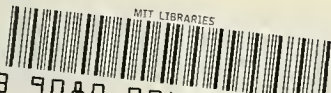
3 9080 004 415 391

MIT LIBRARIES



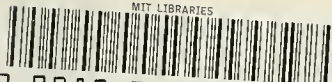
3 9080 004 415 409

MIT LIBRARIES



3 9080 004 415 417

MIT LIBRARIES



3 9080 004 415 425

